Photorealism enhancement (image-to-image translation)

* We are going to investigate the problem of making rendered or synthetic images more realistic. The problem of achieving realistic images is crucial for various applications, such as gaming, where realistic visuals enhance user immersion and engagement, virtual and augmented reality, simulation, digital art and photoraphy, and also scientific or medical visualization.
* This problem has already been highly approached during the past, so we will examine many papers relative to image enhancing and translation such as:

<https://paperswithcode.com/paper/enhancing-photorealism-enhancement>

<https://paperswithcode.com/task/synthetic-to-real-translation>

<https://paperswithcode.com/task/image-to-image-translation>

<https://paperswithcode.com/paper/art2real-unfolding-the-reality-of-artworks>

<https://openaccess.thecvf.com/content_ECCV_2018/papers/Chuanxia_Zheng_T2Net_Synthetic-to-Realistic_Translation_ECCV_2018_paper.pdf>

* As regards the dataset, we will evaluate three options:

1. We can create our own dataset, selecting real photos from Internet and create their corresponding rendered version by applying filters, such as blurring, color adjustements, noise and texture filters, artistic filters and geometric transformations, such as perspective distortion, rotation, …
2. We can use the LAION-400-MILLION dataset, that contains 400 million CLIP-filtered image-text pairs. We would select a subset of real images from LAION and pair them with the relative filtered images throgh the CLIP filtering. We would then preprocess them, using filters and geometrical transformations.
3. We can select synthetic images from a dataset like GTAV-Cityscapes, and use a CNN like ResNet to find and retrieve the similar real images from another dataset, such as Cytiscapes, COCO or OpenImages. We would then do the preprocessing in a similar way to option 2.

* We intend to study the literature to understand what are the methods that give the best and most realt-time results on this kind of task. In particular, we have to decide if it’s better to use a GAN, such as StyleGAN or CycleGAN, or use a DIFFUSION model. In any case, the idea is to use a pretrained network, given the limited amount of computational power at our disposal, finetune it on our dataset and modify some layers/parameters in order to obtain the most real-like images.
* The results will be evaluated in terms of FID and CLIP, to evaluate the quality of the generated images, and in terms of Recall, for what regards the retrieval of real images from the dataset, similar to the generated ones. We can compare our results against the ones obtained in the cited papers by using other types of networks (es another type of GAN).

PIPELINE:

If we choose to create our own dataset, we will use **Classical image-processing operators,** such as Bilateral filter, Posterization filter, Gaussian blur, and **Geometric-based algorithm**, such as Rotation and Perspective Transformation to create the corresponent rendered images to real ones. While if we choose to use an existing dataset, the filters and transformations will be applied to preprocess the images.

The **Retrieval algorithm,** such as Template matching or Feature-based matching like SIFT, will be used to find the real images from the dataset similar to the images generated from our network. Moreover, if we choose option 3 in the dataset choice, a Retrieval algorithm will be used to find the correspondent real images from the synthetic ones of the GTAV dataset.

Lastly, as **Deep learning-based component** we will implement a generative network, such as GAN or DIFFUSION, and finetune it by training the last layers.

Options:

* Image restoration/ super-resolution

<https://arxiv.org/pdf/2308.09388v1>

<https://proceedings.neurips.cc/paper_files/paper/2023/file/2ac2eac5098dba08208807b65c5851cc-Paper-Conference.pdf>

* From sketch to realistic image

<https://arxiv.org/pdf/2208.12675>

<https://github.com/inzva/sketch-to-photograph-with-GANs/blob/master/README.md>

* Improving realism on synthetic data

<https://arxiv.org/pdf/2304.12463>